1. a) $n = 75$, $\alpha = 0.05$, $Z(0.025) = 1.96$
The 95% confidence interval
is $\bar{X} \pm Z^* s/\sqrt{n}$ --> $3433 \pm 1.96^* 495/\sqrt{75}$ --> (3320.971, 3545.029)

b) $n = 75{,}000$
The 95% confidence
interval is $3433 \pm 1.96^* 495/\sqrt{75000}$ --> (3429.457, 3436.543)

c) The confidence interval in a ~~is~~ wider.
The smaller sample results in an
estimate that is not as accurate.

Scanned with CamScanner

Scanned with CamScanner

(2) $n = 36$ population : $\mu = 91.4$

sample : $\bar{x} = 92.8$ $\sigma = 3.6$

$\alpha = 0.05$

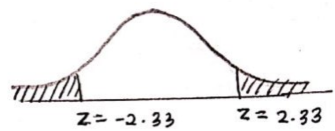
(1) Hypothesis statement :-

$H_0 : \mu = 91.4$

$H_1 : \mu \neq 91.4$

(2) Test statistic :-

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \qquad z = \frac{92.8 - 91.4}{\frac{3.6}{\sqrt{36}}}$$



$$z = 2.3333$$

$P(z \leq 2.3333) = 0.99010$

$P(z > 2.3333) = (1 - 0.99010)$

$= 0.0099$

$\therefore$ p-value $= 2(0.99010) = 1.9802$

since $1.9802 > 0.05$, fail to reject H_0

$\therefore$ Fail to reject H_0, there is sufficient evidence to conclude that $\mu = 91.4$.

Question 3

The claim that male drivers are more prone to go through zebra crossing compare to female drivers.

if $P_1 > P_2$ is false than $P_1 \leq P_2$.

$H_0 : P_1 = P_2$ $\quad \alpha = 0.05$ $\quad n_1 = 7, x_1 = 51$

$H_1 : P_1 > P_2$ $\quad n_2 = 53, x_2 = 31$

$\hat{P}_1 = 51/71$; $\hat{P}_2 = 31/53$

$= 0.7183$ $\quad = 0.5849$

	drivers	no of drivers
Male	$n_1 = 71$	$x_1 = 51$
female	$n_2 = 53$	$x_2 = 31$

$\bar{P} = \dfrac{51 + 31}{71 + 53}$ $\quad \bar{q} = 1 - \bar{P}$

$= 1 - 0.6613$

$= \dfrac{82}{124} = 0.6613$ $\quad = 0.3307$

test statistic $z = \dfrac{0.7183 - 0.5849}{\sqrt{(0.6613)(1-0.6613)(1/71 + 1/53)}}$

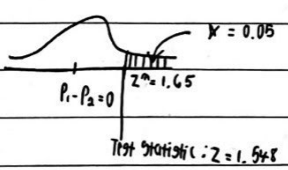
$= 1.5528$

P- value of $z = 1.5528$

Since P - value of $z = 1.5528$ this is a right - tailed test, thus P-value $= 0.0602$.

Since P- value $= 0.0602 > \alpha = 0.05$, therefore we fail to reject the null hypothesis, H_0.

$\therefore$ There is insufficient evidence to support the claim that male drivers are more prone to go through zebra crossing compare to female drivers.



4) Test : $\mu 1 \neq \mu 2$ using ~~two~~ sample t-test

Hypothesis test

$H_0 : \mu 1 - \mu 2 = 0$ (Null Hypotesis)

$H_a : \mu 1 - \mu 2 \neq 0$ (Alternate Hypothesis also called H_1)

This is ~~two~~ tailed test

sample 1 (strawberry flavor)	sample 2 (blueberry flavor)
$\bar{x}_1 = 21.033333$	$\bar{x}_2 = 20.892308$
$s_1 = 0.606218$	$s_2 = 1.007027$
$n_1 = 9$	$n_2 = 13$

Significance level, $a = 0.05$

Since variances are not equal,

Degrees of freedom $= df =$ smaller of $n_1 - 1$ or $n_2 - 1 = 8$

Standard Error : $\sigma \bar{x}_1 - \bar{x}_2 = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{0.3675/9 + 1.014103/13} =$

$0.344734 \approx 0.3447$

Test statistic, t observed $= \frac{\bar{x}1 - \bar{x}2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$

$t_0 = \frac{21.033333 - 20.892308}{\sqrt{0.3675/9 + 1.014103/13}} = 0.409084 \approx 0.409$

Using excel we find p-value $= 2 * P(t > |0.409|) = 2 *$ t.dist(0.409, 8, 1

$= 0.69321084 71$

Note that some books teach Welch's t-test, where approximate degrees of freedom are :

$$v \approx \frac{\left(\frac{s_1^2}{n1} + \frac{s_2^2}{n2}\right)^2}{\frac{s_1^4}{n_1^2(n_1-1)} + \frac{s_2^4}{n_2^2(n2-1)}}$$

Scanned with CamScanner

For welch t-test, p-value is 0.68705 25145
And t-critical would have been: ± 2.093

Critical value t critical is give by $t_{a/2, df=8} = 2.306$ (from t-table, one-tail, d.f. -8)

$t_c = \pm 2.306$

Rejection criteria: Rejection region is in direction of alternate hypothesis; since $H_a: \mu_1 - \mu_2 \neq 0$, rejection rule is:
~~since~~ Reject H_0 if $|t_a| > t_c$ or equivalently if p-value $< a$

Decision: $|t_a| \leq t_c$ or equivalently since p-value $\geq a$, we fail to reject the null hypothesis. There is insuffcient evidence to conclude that μ_1 is different from μ_2

Scanned with CamScanner

5

$n = 15$

$\alpha = 0.05$

2 hour

$\bar{x}_1 = \cancel{3921.93}$ 3924.0667

$s_1^2 =$ 688300.78

4 hour

$\bar{x} = 4069.266667$

$s_2^2 = 908010.5$

Hypothesis statement:

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

Given $\alpha = 0.05$. Test statistic is $t_0 = \dfrac{\bar{x}_1 - \bar{x}_2 - 0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$

~~reject H_0 if $t_0 >$~~

$$s^2_p = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{(14)(688300.78) + (14)(908010.5)}{14 + 14 - 2}$$

$$= 1596323.28$$

$$s_p = \sqrt{1596323.28} = 1263.456877$$

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2 - 0}{1263.456877\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{3924.0667 - 4069.266667}{1263.4569\sqrt{\frac{1}{15} + \frac{1}{15}}}$$

$$= -0.3147$$

$\therefore$ since $-2.145 < t_0 = -0.3147 < 2.145$, fail to reject H_0,

That is at the 0.05 level of significance, we do not have strong evidence to conclude that differ mean of food intake for group 2 hour and 4 hour playing e-sport.

Date:

Question 6

$H_0 = \mu_0 = 3$ and $H_a = \mu_0 > 3$ (right-tailed test)

$df = 8 - 1 = 7$; $n = 8$

$\alpha = 0.1$

Before Treatment (X_1)	After Treatment (X_2)	$D = X_1 - X_2$	$D^2 = (X_1 - X_2)^2$
14	10	4	16
12	4	8	64
18	14	4	16
7	6	1	1
11	9	2	4
9	6	3	9
16	12	4	16
5	12	3	9
		$\Sigma D = 29$	$\Sigma D^2 = 135$

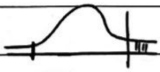
$$\bar{D} = \frac{\Sigma D}{n} = \frac{29}{8} = 3.625$$

$$S_D = \sqrt{\frac{\Sigma D^2 - \frac{(\Sigma D)^2}{n}}{n-1}} = \sqrt{\frac{135 - \frac{(29)^2}{8}}{8-1}} = \sqrt{\frac{29.875}{7}} = 2.066$$

$$t = \frac{\bar{D} - \mu_0}{S_D / \sqrt{n}} = \frac{3.625 - 3}{2.066 / \sqrt{8}} = 0.856$$

Decision : Since, 1.415 > 0.856, the decision is to reject the null hypothesis at

Conclusion : Hence, there is enough evidence to support the claim that the new teaching method is possible to increase memory recall for slow learners.



0.856 1.415

7. $N = 1194$

$\alpha = 0.05$

$n(\text{size}) = 7$; $n \leq 30$

$H_0 = N = 1194$

$H_1 = N \neq 1194$

$\mu = \frac{1194}{7}$

$= 170.57$

$\bar{x} = 1194$

$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$

$= \frac{1194 - 170.57}{451.29/\sqrt{7}}$

$= \frac{1023.43}{170.572}$

$= 6.0$

$s = \sqrt{\frac{n \Sigma x^2 - (\Sigma x)^2}{n(n-1)}}$

$= \sqrt{\frac{7(1194)^2 - (1194)^2}{7(6)}}$

$= \sqrt{203662.29}$

$= 451.28958$

$= 451.29$

$c.v: t_{0.05, 1194} = 6.00$

Fail to reject H_0: There is sufficient evidence that not support the standard deviation in the length of convincing evidence of male smoker lungs.

Scanned with CamScanner

8) yellow and sweet - 124
green and sweet - 30
yellow and juicy - 43
yellow and sour - 11

$\alpha = 0.01$ ratio 9 : 3 : 3 : 1

observed	ratio	expected	$\frac{(O-E)^2}{E}$
124	9	117	0.42
30	3	39	2.08
43	3	39	0.41
11	1	13	0.31
208	16	208	3.21

chi-square value critical value χ^2 : 11.345

: $\chi^2 = \frac{\Sigma(O-E))}{E} = 3.21$

- We do not reject because test statistic value ($\chi^2 = 3.21$) < critical value (11.345) as it shows it is not fall within critical region.

Scanned with CamScanner

9. H_0 = group assignment is independent.
H_1 = group assignment is not independent.

Salary (Thousand of dollars)

Gp	27-29		29-31		31-33		33-35		≥ 35		Total
	D	E	D	E	D	E	D	E	D	E	
1	6	7.33	11	13.33	16	16	14	13.33	13	10	60
2	5	3.67	9	6.67	8	8	6	6.07	2	5	30

cell, ij	Observed count, O_{ij}	Expected count, E_{ij}	$[O_{ij} - E_{ij}]^2/E_{ij}$
1,1	6	11×60/90 = 7.33	0.2413
1,2	5	11×30/90 = 3.67	0.4820
2,1	11	20×60/90 = 13.33	0.4073
2,2	9	20×30/90 = 6.67	0.8139
3,1	16	24×60/90 = 16	0.0000
3,2	8	24×30/90 = 8	0.0000
4,1	14	20×60/90 = 13.33	0.0337
4,2	6	20×30/90 = 6.67	0.0673
5,1	13	15×60/90 = 10	0.900
5,2	2	15×30/90 = 5	1.800
			$\chi^2 = 4.7455$

Critical value:

$v = (5-1)(2-1) = 4$

$\alpha = 0.10$

$\chi^2_{0.10,4} = 7.779$

since $\chi^2 = 4.7455 < \chi^2_{0.10,4} = 7.779$

Fail to reject H_0: There is sufficient evidence that support the group assignment is independent.

Scanned with CamScanner